DOCUMENT RESUME

ED 363 977 EA 025 507

AUTHOR Wheeler, Patricia; And Others
TITLE Teacher Evaluation Glossary.
INSTITUTION Center for Research in Educational Accountability and Teacher Evaluation (CREATE), Kalamazoo, MI.
SPONS AGENCY Office of Educational Research and Improvement (ED), Washington, DC.
PUB DATE Aug 92
CONTRACT R117Q00047
NOTE 35p.
AVAILABLE FROM Education Center, Western Michigan University, Kalamazoo, MI 49008-5178 ($3).
PUB TYPE Reference Materials - Vocabularies/Classifications/Dictionaries (134)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Accountability; Elementary Secondary Education; Evaluation Methods; Glossaries; Instructional Improvement; Performance; *Teacher Competencies; *Teacher Evaluation; Vocabulary

ABSTRACT

This glossary addresses the increasingly complex vocabulary of teacher evaluation. It lists 265 terms that are frequently used in the process of evaluating practicing teachers, particularly in elementary and secondary schools, and provides brief explanations of the meanings of these terms. This glossary is a product of one of the several projects supported by the Center for Research on Educational Accountability and Teacher Evaluation (CREATE)--the Teacher Evaluation Models Project. It is hoped that this glossary will help to improve teacher evaluation through a shared understanding of key terms. Sources of additional information on the terms--professional organizations and references--are listed at the end of the document. (LMI)

**********
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
**********

ED 363 977

# TEACHER EVALUATION GLOSSARY

**Patricia Wheeler, Ph.D.**
**Geneva D. Haertel, Ph.D.**
**Michael Scriven, D.Phil.**

**January 1993**

**Center for Research on Educational Accountability and Teacher Evaluation**
**The Evaluation Center**

**WESTERN MICHIGAN UNIVERSITY**

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
P. Wheeler

BEST COPY AVAILABLE

2 TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

EA 025 507

**For more information contact:**

**Dissemination Officer**
**CREATE**
**The Evaluation Center**
**Western Michigan University**
**Kalamazoo, MI 49008-5178**
**Phone: (616) 387-5895**
**Fax: (616) 387-5923**
**Internet E-Mail CREATE_Info**

**This work was supported by the Office of Research, Office of Educational Research and Improvement, U.S. Department of Education (Grant No. R117Q00047). The opinions expressed are those of the authors, and no official support by the U.S. Department of Education is intended or should be inferred.**



BEST COPY AVAILABLE

# TEACHER EVALUATION GLOSSARY

Patricia Wheeler, Ph.D.
Geneva D. Haertel, Ph.D.
Michael Scriven, D. Phil.

A Product of the Teacher Evaluation Models Project,
Center for Research on Educational Accountability & Teacher Evaluation,
The Evaluation Center,
Western Michigan University
Kalamazoo, Michigan

August 1992

## *Teacher Evaluation Models Project Advisory Board*

*David Berliner*
*Arizona State University*

*Edwin Bridges*
*Stanford University*

*Linda Darling-Hammond*
*Teachers College*
*Columbia University*

*Gene Glass*
*Arizona State University*

*Jason Millman*
*Cornell University*

*Kenneth Peterson*
*Portland State University*

*W. James Popham*
*University of California at Los Angeles*

*Lee Shulman*
*Stanford University*

## *R&D Center Staff*

*Daniel Stufflebeam, CREATE Director*

*Arlen R. Gullickson, CREATE Associate Director*

## *Project Staff*

*Michael Scriven, TEMP Project Director*

*Patricia Wheeler, TEMP Senior Research Associate*

*Geneva D. Haertel, TEMP Research Associate*

# PREFACE

This glossary addresses the increasingly complex vocabulary of teacher evaluation. It lists 265 terms that are frequently used in the process of evaluating practicing teachers, particularly in elementary and secondary schools, and provides brief explanations of the meanings of these terms. The explanations are based upon the references listed in the back of this document, as well as on the professional experiences of the authors, of the advisory committee for this project, and of other staff at the Center for Research on Educational Accountability and Teacher Evaluation (CREATE) at Western Michigan University. CREATE is one of several research centers across the United States that focus on different issues in education and are funded by the Office of Educational Research and Improvement (OERI) of the U.S. Department of Education. This glossary is a product of one of the several projects supported by CREATE—the Teacher Evaluation Models Project (TEMP), and one of its functions is to provide definitions to help readers of other TEMP publications, particularly TEMP's series of memoranda on key issues in teacher evaluation. Readers interested in being on the (free) mailing list for these memoranda can request them from CREATE, Evaluation Center, Western Michigan University, Kalamazoo, MI 49008-5178. Copies of this glossary may be obtained from the same address for $3.00 including postage and handling.

The realization that a glossary like this could be useful arose from our work with teachers and administrators at schools we visit or who attend workshops on teacher evaluation, or who write to us about it. It is clear that there is a wide variation in the use of many terms from one part of the country to another, and from one country to another, especially because of the differences between the models of teacher evaluation standardized by State Departments of Education.

It is hoped that this glossary will make a small contribution to improved teacher evaluation through shared understanding of key terms and concepts by the various people involved in the process of teacher evaluation and in discussions of the process. The authors welcome suggestions for changes and additions; please send them to Patricia Wheeler, 2840 Waverley Way, Livermore, CA 94550-1740.

The glossary forms the third and smallest of the three major "infrastructure publications" of TEMP, aimed to facilitate future research and development in the field of teacher evaluation, by TEMP itself and by others. The other two are a substantial *selective* bibliography of the relevant literature (about three thousand items) and a set of abstracts of what we judge to be the most significant subset of that literature (about six hundred items). Both are considerably more comprehensive and up to date than anything else available. These two projects are unlikely to be published in hard copy form because of their size (over 600 pages) and limited market but the latest versions of both can be obtained from

CREATE on a single disk for the cost of copying, postage, and handling ($6). The standard format is Microsoft Word on a Macintosh 1.4 MB disk, which can be readily converted into MS-DOS format and any word processor format at many copy shops and any conversion service. By the beginning of 1993, we expect a more finished version to be available through a commercial publisher in several formats, probably as an optional accessory to a book containing much of the TEMP materials including this glossary.

The entries here include some relatively new terms as well as many that are well established. Some terms have meanings that vary from context to context, and are here defined in the sense that is specific to the teacher evaluation context. There are also general definitions of a few key terms from measurement and evaluation that are widely used in discussions of teacher evaluation. Terms were selected for frequency of use in the field, or for their importance in understanding key concepts directly related to teacher evaluation.

Since this is a glossary and not a thesaurus or encyclopedia, the explanations are brief and have been written using neutral language so as to minimize any appearance of bias toward a particular approach or perspective on teacher evaluation. Note that this is not a glossary of terms used in *supervision*, and in fact there is almost no overlap with a glossary covering 100 terms used in that field[1].

Sources of additional information on the terms in this glossary are listed at the end of this document. They include the addresses of several professional associations and agencies that provide materials on teacher evaluation, as well as the references used in preparing this glossary.

In addition to the Advisory Board, who saw an early draft, the following people provided particularly valuable help in the development of this document: Ruben Abrica, Iris Berke, Eugene Ellis, Geraldine Erickson, Edward Haertel, Andrea Lash, Virginia Luhring, Jean Martinson, Jason Millman, Larry Murphy, Ray Novak, Michael Reed, Stanley Rose, Joanne Savage, William H. Schubert, James Stronge, Don Ulrey, and William Wiersma.

The preparation of this document was supported under the Educational Research and Development Center Program Grant Award Number R117000047 as administered by the Office of Educational Research and Improvement (OERI) of the U. S. Department of Education (USED). The findings and opinions expressed in this document do not necessarily reflect the position or policies of OERI or USED.

Patricia Wheeler
Geneva D. Haertel
Michael Scriven
*August 1992*

---

[1] Dzyacky, John W., (Ed.) (1987). *Glossary: 100+1 Terms to know and apply in supervising instruction.* Ames, IA: The School Improvement Model Project, Iowa State University.

**Ability** – the capacity of a teacher to perform a task or to use intellectual or physical skills. Contrasted with *performance* (the use of abilities) and *aptitude* (the capacity to acquire abilities). *See* Aptitude, Performance, and Skills.

**Accountability** – the responsibility for implementing a process or procedure, for justifying decisions made, and/or for results or outcomes produced. Teachers are often said to be accountable for their students' learning in the assigned subject area, within the limits of the students' abilities and the time and resources available. However, there are some limits on the extent to which the teacher can be held responsible for motivation of students and hence the part of learning dependent on it.

**Accuracy** – the degree to which the data and information collected about the performance of a teacher are precise and correct measures of performance and are free from error. One of the Personnel Evaluation Standards. *See* Bias, Conflict of interest, Credible, Error of measurement, Rater effect, Verification.

**Administrative feasibility (of teacher evaluation)** – the extent to which appropriate data are readily available or can be obtained, produced or interpreted with available—or, sometimes, obtainable—resources such as staff expertise, time, and equipment.

**Administrative responsibility (for teacher evaluation)** – accountability, often stated in the teacher evaluation policy, for the conduct of the evaluation process, and decisions based on it.

**Administrator** – a manager in the organization within which the teacher works; one manager is normally in charge of the evaluation of teaching, and often also responsible for ensuring that teachers have, to the extent possible, the resources needed to perform their duties and professional responsibilities. *See* Supervisor.

**Adverse impact** – a difference between certain identified groups in the effect of evaluation results on personnel decisions. The Equal Employment Opportunity Commission's indicator for possible adverse impact is that the selection rate for any race, sex, or ethnic group is less than 80 percent of the rate for the group with the highest rate. Comparable qualifications should be ensured.

**Affective domain** – the area of feelings and emotions including interests, attitudes, motivations, values, and appreciations. *See* Cognitive domain, Psychomotor domain.

**Analysis** – the treatment of data and information in order to elicit certain statistical results and evaluative conclusions. *See* Congruence analysis, Quality check, Synthesis.

**Analytic scoring** – an approach to scoring or rating that considers various parts or aspects of the attribute or performance being assessed, for use in profiling strengths and weaknesses or in obtaining an overall summary. Scores may be recorded as a check mark for presence or absence of an attribute, marked on a rating scale, or put in the form of a brief comment. *See* Holistic scoring, Scoring, Scoring rubric.

**Anecdotal record** – a short narrative of an event or activity that may be used to support generalizations about the performance of a teacher.

**Announced observation or visit** – one that is prearranged with the teacher who is to be evaluated, and for which the teacher can prepare.

**Appeal process** – a procedure by which the teacher or another stakeholder can challenge the results of an evaluation of a teacher. An appeal may lead to a formal hearing. *See* Grievance, Hearing, Stakeholders.

**Aptitude** – Capacity for developing skills, knowledge, or attitudes; roughly, potential for acquiring an ability. *See* Ability.

**Assessment** – the process of measuring, quantifying, and/or describing those aspects of teaching related to an evaluation; sometimes, the overall process of teacher evaluation.

**Assessment center** – literally, a place where teaching performance is measured using judges observing simulations. The term is sometimes used as a name for an approach to teacher evaluation, without reference to a specific location.

**Assessor** – the person who collects data and who measures attributes related to the performance of a teacher; the term has much the same meaning as evaluator. Assessors may be principals, other teachers, students, parents, district staff, or other persons. This term includes, but is not limited to, the interviewer, judge, observer, and scorer. *See* Evaluator, Interviewer, Judge, Observer, Scorer.

**Attribute** – a characteristic, behavior, skill, type of knowledge, or perceived quality such as, but not limited to, attitude, ability, or interest. *See* Construct.

**Audience** – those individuals who have a potential interest in the results of teacher evaluation. *See* Stakeholders.

**Audit** – an independent quality check and verification of the evaluation of a teacher. *See* Quality check, Verification.

**Authentic assessment** – an assessment approach that has been designed to provide a realistic task, simulation, or problem related to that attribute or performance being measured. *See* Portfolio, Simulation.

**Bargaining unit** – *See* Collective bargaining unit.

**Behaviors** – the actions of the teacher or others, including the students, that are specific and observable.

**Bias** – (1) a systematic tendency to a lack of objectivity, fairness, or impartiality on the part of the assessor or evaluator, often based on his or her personal preferences and inclinations; (2) systematic error in the assessment instrument or evaluation process. *See* Contamination, Error of measurement, Rater effect.

**Career ladder** – an incremental scale of professional teaching positions through which a teacher advances, based on evaluations of past performance.

**Central tendency effect** – a type of rater effect in which an assessor or an evaluator tends to rate teachers toward the mid-point of a scale, or to judge the performance as average or of intermediate merit, when it is actually well above or well below the middle level of the scale. (Note: this is a misleading use of a term from descriptive statistics; the terms "central bias" or "centrism" might be less confusing.) *See* Rater effect, Leniency, Stringency.

**Certification** – (1) official recognition of advanced status, outstanding performance or high level of expertise in the profession, usually granted to incumbent

teachers who have several years of teaching experience; (2) sometimes used as a synonym for credential or license. *See* Credentialing, License.

**Checklist** – an instrument that specifies criteria or indicators of merit on which the assessor or evaluator marks the presence or absence of the attribute being assessed. Its use can improve inter-rater and intra-rater reliability.

**Client** – the agency or individual who commissioned the evaluation and to whom the evaluator has legal responsibility.

**Clinical supervision** – a process of collaboration between the teacher and the supervisor or administrator, designed to improve the teacher's performance. This process usually includes pre-observation conference, observation and data collection, data analysis, post-observation conference, and an evaluation report following the post-observation conference.

**Coaching** – assistance to teachers in ways that will improve their job performance, not just for the purpose of doing better on the evaluation, but also for the purpose of improved teaching and increased student learning. Coaching can also refer to teachers coaching students so that the students will perform better on a measure used as an indicator of the teacher's performance. *See* Peer review, Test pollution.

**Code** – a symbol, either numeric or alphabetic, used to represent attributes, words, or behavior (e.g., 3 = Third Grade, M = Math, TQ = Teacher asks question).

**Coding** – the process of assigning codes to data, evidence, information, judgments, notes, and responses. *See* Code.

**Cognitive domain** – the range of knowledge and knowledge-related skills, ranging from perceptual skills and recall of facts to complex reasoning. *See* Affective domain, Psychomotor domain.

**Collective bargaining agreement** – a written document, approved by representatives of administrators, policymakers and teachers, that addresses concerns such as teacher salaries, benefits, working conditions, evaluations, terminations and dismissals, appeal procedures, fair hearings, teachers' rights, and other aspects of teaching. *See* Contract, Hearing, Teachers' rights.

**Collective bargaining unit** – a group that represents the teachers' interests in negotiations with administrators and policymakers.

**Comments** – the information provided by anyone involved in the evaluation process (e.g., teacher being evaluated, observer, interviewer, data collector, judge, analyzer, evaluator) concerning incidents or factors that could affect the quality and accuracy of the assessment data and the judgments and evaluations made (e.g., fire alarm sounded during the pre-conference interview).

**Comparability** – the essential similarity of phenomena (e.g., attributes, performances, assessments, data sources) being examined. The amount or degree of comparability is often used to determine the appropriateness of using one phenomenon in lieu of another; and often to ensure fairness. *See* Equivalent.

**Compensatory model** – an evaluation procedure or scoring procedure that permits trade-offs between weakness in one attribute and strength in another. Usually, there is a minimum required level of achievement on several of the at-

tributes, so this is an overlay on a conjunctive model, the idea being that 'overscores' can be combined to meet some further requirement—e.g., a minimum total score—or in order to get an overall score for purposes of ranking. If there are no minimums, this is a quantitative variety of the disjunctive model, i.e., evaluees must acquire a certain number of total points, but can do so even if scoring very poorly on some dimensions. Teacher evaluation excludes this kind of simple compensatory approach with no minimums, a point of view sometimes expressed in language like this: "good knowledge of subject matter is no substitute for being able to communicate with the youngsters in your class, nor vice versa." There are also hybrid models with *absolute* minimums on some dimensions, and a requirement of also achieving a higher-than-absolute-minimum score on a certain number of dimensions, beyond which point an individual's scores are simply pooled (i.e., strong scores on any dimension are allowed to compensate for weak scores on others). *See* Conjunctive model, Disjunctive model.

**Competence (teacher)** – a teacher's repertoire of competencies. *See* Competency.

**Competency (teaching)** – a knowledge, skill, ability, personal quality, experience or other characteristic that is applicable to the profession of teaching. *See* Competence.

**Component** – (1) one of the parts or processes in an evaluation system (e.g., pre-observation conference, group interview, classroom observation, portfolio); (2) one of the parts of an individual teacher's performance that is assessed in order to make up the total appraisal of that teacher. Components are distinguished from dimensions in that they are separate in space or time. *See* Dimension.

**Composite score** – a score that combines two or more scores or results for different attributes or dimensions of performance. It is sometimes necessary to calculate composite scores, e.g., in terms of rank candidates, but it is to be avoided if possible since justification of the combination rule is often difficult.

**Computerized assessment** – the use of computers to measure performance on some attribute, not necessarily an attribute related to computers and technology. Now widely used for student assessment; increasingly relevant for teacher assessment, e.g., of subject matter knowledge, of pedagogical knowledge.

**Concurrent validity** – *See* Validity

**Conference** – a meeting between the teacher and the assessor or evaluator to discuss mutual concerns and to promote understanding of the assessments being used and the evaluation process and results. The conference can also be an opportunity to collect teacher responses if the conference includes an interview. *See* Interview, Responses.

**Confidentiality** – the protection of data and information from persons other than those authorized to have access. Authorized access may be possible for people who have not been specifically 'cleared' for access by the individual to whom the data refers.

**Conflict of interest** – a situation in which the private interests of someone involved in the evaluation process (e.g., interviewer, scorer, evaluator) have potential impact (either positive or negative) on the quality of the evaluation activities, the accuracy of the data, or the results of the evaluation. *See* Accuracy.

**Congruence analysis** – the verification of data by using more than one source of data for assessing performance on the same criterion. *See* Triangulation.

**Conjunctive model** – an evaluation or scoring procedure that requires the teacher to attain a minimal level of performance on all attributes assessed. Sometimes referred to, perhaps less obscurely, as the multiple-cutoff model. *See* Compensatory model, Disjunctive model.

**Considered necessary** – that which is judged to be required, but not sufficient.

**Construct** – an attribute of an individual that is not directly observable but is inferred from empirical evidence; e.g., being a 'reflective teacher.' *See* Attribute, Validity (Construct).

**Construct validity** – *See* Validity.

**Contamination** – a tendency for the assessor's judgments or the evaluator's conclusions to be influenced by irrelevant knowledge about the teacher, other personnel, or other factors that have no bearing on the teacher's level of performance. *See* Bias, Error of measurement, Rater effect.

**Content validity** – *See* Validity.

**Context (teaching)** – the environment within which the teacher works. This includes, but is not limited to, physical facilities and setting, types of students, school and community characteristics, resource availability (staff, materials, equipment, funding, time), classroom climate, school climate, degree of support provided by others, and demands made on the teacher. *See* Learning environment.

**Contextual variables** – *See* Notes.

**Contract** – an agreement (written or oral) between two or more parties concerning expectations and responsibilities of the parties. *See* Collective bargaining agreement, Plan of assistance.

**Contrast effect** – a type of rater effect in which an assessor or an evaluator tends to compare one teacher to other teachers rather than comparing that teacher's level of performance to the standards. *See* Rater effect.

**Correlate** – a factor which is statistically correlated with (usually) effective teaching. *See* Effective teaching.

**Cost** – the quantity of resources required in order to achieve a desired end. The cost of teacher evaluation involves such factors as time, energy, and stress that are not completely reducible to money terms.

**Credentialing** – the process of reviewing potential teachers' qualifications and issuing licenses to teach. *See* Certification, License.

**Credible** – worthy of confidence and acceptance by others, usually based on the expertise, trustworthiness and/or reliability of the source of the evidence or judgment. Credible does not necessarily mean accurate or valid. *See* Accuracy, Reliability, Trust, Validity.

**Criterion, Criteria** – a dimension on which performance (e.g., effective teaching) is rated or judged as successful or meritorious. Each criterion falls within a **domain** covered by the evaluation system and, in one terminology, is defined by **elements, indicators**, and **descriptors** (see below for examples). Indicators and

descriptors are supposed to be stated in measurable or observable terms. Satisfactory levels of performance on criteria are specified by standards. *See* Foundation, Standard.

Example of Criterion—*Teacher can select and create materials that are related to the subject area and are appropriate for the students.*

**Domain** – a broad area covered by a teacher evaluation system and for which criteria and standards are specified for assessing performances in that domain.

Example of Domain – *I. Knowledge of Instructional Design*

**Element** – a major category of teacher knowledge, skills, abilities, and behaviors within a domain.

Examples of Elements for Domain I. – *I.A. Planning of Courses and Lessons; I.B. Selection and Creation of Instructional Materials*

**Indicator** – for each element, the types of knowledge, skills, abilities, behaviors, and attributes that are empirically or by definition connected to the criterion.

Examples of Indicators for Element I.B. – *I.B.1. Materials selected/created fit into instructional plan. I.B.2 Materials selected/created are current, correct, comprehensive. I.B.3. Materials created by the teacher are readable to the students in terms of level of content difficulty, design and printing quality.*

**Descriptor** – for each indicator, a specific example of the performance being assessed, or an observable element of the indicator.

Examples of Descriptors for Indicator I.B.1. – *I.B.1.a. Teacher's handout refers to some topics covered in previous lessons. I.B.1.b Homework assignment addresses two of the district's instructional goals and objectives in this subject area for this grade level.*

**Criterion-related validity** – *See* Validity.

**Critical incident** – a significant event in a teacher's career that alters future teaching performance or events. Such events may include conspicuous successes or failures that are not typical of the teacher's performance, but which should be considered as potential information for the evaluation of that teacher's performance.

**Critical score** – *See* Cutting score, Passing score.

**Curriculum** – (1) a comprehensive overview of a particular subject matter or performance area, including activities planned for delivery to the students, the scope of content, the sequence of materials, interpretation and balance of subject matter, and motivational, instructional, and assessment techniques to be used; (2) sometimes, a set of ordered, intended, learning outcomes covering such an area.

**Cutting score** – a score or level of performance that marks the difference between two types of teachers (e.g., good and excellent). When the difference is between minimally acceptable and not acceptable, or pass and not pass, it is sometimes referred to as a cut score, critical score, or passing score. *See* Passing score.

**"Dance of the lemons"** – the practice of reassigning teachers who are incompetent or who are performing below acceptable levels to other positions in the school or the district (rather than proceeding against them). This practice is also called "pass the turkey" or "turkey trot." *See* Incompetence.

**Data** – the items and observations gathered during the evaluation process for use in determining the level of teaching performance.

**Data collection procedures** – the steps and sources used to obtain qualitative and quantitative information about a teacher's qualifications and performance. *See* Procedures (evaluation).

**Data integration** – the merging of related data for use in scoring, judging and evaluating.

**Data sources** – the persons, documents, products, activities, events, and records from which the data are obtained.

**Decision rules** – the guidelines for determining the level of merit, value, or worth of an aspect of a teacher's performance from the data; or for connecting those levels with personnel actions such as tenure.

**Defensible** – a conclusion or statement that is explainable or justifiable, based on a solid foundation and policy, explicit evaluation procedures, valid assessments, and sound evaluation practices.

**Definition** – the key reference that specifies the meaning of some element of the evaluation process (e.g., attribute, domain, standard, instrument, assessor); it may be or include a description, explanation, exemplification, contrast, or interpretation, instead of or as well as a conventional definition.

**Descriptor** – *See* Criterion.

**Diagnosis** – the process of determining a teacher's strengths and weaknesses, based on the results of an evaluation. Diagnosis is an essential preliminary to preparing a professional development plan for a teacher or a plan for assistance in cases where remediation is needed. *See* Plan of assistance, Remediation.

**Differential functioning** – a characteristic of an assessment approach, instrument, or evaluation system, that yields higher results for one group than another group, even though both groups have the same level of ability or competence on that aspect of teaching. A form of invalidity.

**Dimension** – An aspect or attribute of a teacher's achievements or of a system of teacher evaluation. *See* Attribute, Component.

**Disjunctive model** – an evaluation or scoring procedure that only requires the teacher to achieve a minimal level of performance on one of the components or dimensions assessed, or measures used. This model is defensible for performance within such domains as 'contribution to the profession,' where there are several alternative ways to achieve an acceptable contribution. It is not appropriate for combining performances on all the domains relevant to satisfactory teacher performance, many of which must be passed separately, hence requiring the conjunctive approach. It is the model used for dealing with the scores from multiple attempts at a licensure test, where a pass on a retake is treated as acceptable. *See* Compensatory model, Conjunctive model.

**Dismissal** – the involuntary termination of employment which can be based upon a teacher's level of performance, school staffing needs, or reduced resources. *See* Forced resignation, Hearing, Incompetence, Tenure.

**Documentation** – the collection or compilation of all tangible materials including evaluation forms and notes used in the evaluation of a teacher. *See* Record.

**Documenting** – providing tangible evidence and information, in writing or other recorded formats, about the performance of a teacher.

**Domain** – *See* Criterion.

**Drift** – the tendency for assessors and evaluators to unintentionally redefine criteria and standards over time or across a series of ratings.

**Due process** – a teacher's right to fair and impartial treatment as guaranteed by the U. S. Constitution and Bill of Rights, and by various laws (e.g., the Civil Rights Act of 1964) and by related procedural requirements. *See* Dismissal, Hearing, Tenure.

**Duties-based evaluation** – a teacher evaluation approach that is based on what a teacher is legally and professionally required to do as a teacher. Duties and professional responsibilities are specified—to a limited degree—in state laws and regulations, school district policies, job descriptions, and in the normal expectations and demands for a given teaching context. *See* Context (teaching), Duty, Foundation, Responsibility, Teacher norms.

**Duty** – that which a teacher is legally required and morally obligated to do (or not do) as part of his or her job. *See* Duties-based evaluation, Responsibility.

**Effective teaching** – those teaching practices that lead to desirable results such as student learning, typically but not necessarily measured by standardized tests. Often such practices are identified based on correlational research which is referred to as 'process-product research' although it does not in fact establish a cause-and-effect link between particular teaching practices and student learning, especially if the practice is added to a teacher's repertoire. *See* Teacher effectiveness.

**Effectiveness** – an attribute of those schools, teachers, programs, and approaches that meet the needs of students and their society. *See* Stakeholders.

**Efficiency** – an attribute of those schools, teachers, programs, and approaches that balance effectiveness against consideration of the cost of effectiveness. *See* Effectiveness.

**Element** – *See* Criterion.

**Environment** – *See* Context (Teaching), Learning environment.

**Equitable** – that which is fair, impartial and just, e.g., by providing equal opportunity for all.

**Equivalence** – comparability of two or more parallel measures that have been designed to assess the same aspect of teaching (e.g., two different social studies textbook chapters to be analyzed as part of a semi-structured interview; two essay questions on teaching the same content area in math, but to different types of student groups). Should normally be established empirically rather than just judgmentally. *See* Comparability.

**Error of measurement** – the difference between an individual's obtained score and his or her 'true score' on that test, where 'true score' is a technical term defined below. Errors of measurement are due to factors beyond the control of the individual being assessed, and include lack of reliability in the assessment instrument or process, variability of settings for the assessment, limited sampling of teacher performance, bias of the assessor, rater effects, and interactions among such factors. *See* Assessor, Bias, Contamination, Rater effect, Reliability, Sampling of performance, True score.

**Evaluand** – that which is being evaluated (e.g., program, personnel, product, policy, proposal, procedure). *See* Evaluator, Evaluee.

**Evaluation** – the systematic process of determining the merit, value or worth of someone (the evaluee, such as a teacher, student, or employee) or something (e.g., an aspect of their performance, an evaluation policy or procedure). *See* Teacher evaluation.

**Evaluation stages** – the major steps in the teacher evaluation process (e.g., orientation meeting; distribution of copies of job descriptions, evaluation procedures and time lines: scheduling of first conference and first formal observation).

**Evaluation system** – *See* System (Evaluation).

**Evaluator** – in a system of teacher evaluation, *the* evaluator is the person who assembles information collected about a teacher, analyzes it, makes judgments as to whether that teacher's overall performance level meets the pre-specified standards, prepares a summary report, writes recommendations, and may or may not provide feedback to the teacher, directly or through another person. In general, the evaluator is the person who determines the *overall* merit, value, or worth of the evaluee. The term also applies more generally to those supplying subevaluations on particular dimensions (e.g., knowledge of subject matter) that are assembled by the overall evaluator. *See* Assessor, Evaluand, Evaluee.

**Evaluee, Evaluatee** – a person whose qualifications and performance are being evaluated. *See* Evaluand, Evaluator.

**Evidence** – the documentation and verbal statements by credible witnesses (e.g., students, peers, principal) concerning the knowledge, skills, abilities, and behaviors of a teacher. Evidence is used both to generate and to justify judgments about a teacher's performance for purposes of evaluation. *See* Credible.

**Exemplary teacher** – a teacher whose level on some aspect of performance is regarded as deserving of imitation and modeling. *See* Modeling.

**Face validity** – *See* Validity.

**Feasibility** – *See* Administrative feasibility.

**Feedback** – the information and recommendations provided to a teacher about his or her performance based on the results of that teacher's evaluation and designed to help the teacher improve his or her performance.

**First-impression effect** – a rater effect in which an assessor or an evaluator tends to overweight early impressions in forming an overall judgment or a judgement of a later aspect of the performance. *See* Halo effect, Rater effect.

**Forced resignation** – a termination of employment in which the teacher leaves against his or her will, but which is recorded as voluntary (e.g., early retirement instead of layoff). *See* Dismissal, Tenure.

**Formal (evaluation)** – a (formative or summative) evaluation conducted in accordance with a prescribed plan, structure, or advance notice. Also, an evaluation which is pre-identified as determinative of personnel status, although it may be informal in procedure; in this sense, the term means essentially "an official summative evaluation." *See* Informal.

**Format** – the structure of assessment instruments, evaluation forms and materials. Format includes shape, size and general design or layout of the materials.

**Formative teacher evaluation** – an evaluation conducted primarily for the purpose of professional development, i.e., improving the teacher by identifying that teacher's strengths and weaknesses. Most often done by another teacher (or teachers) or by a supervisor. *See* Summative teacher evaluation.

**Foundation** – the basis for the attributes and domains covered by the teacher evaluation. Foundations can include teacher duties and professional responsibilities, research on teaching, governmental policies, professional expertise, and theories of teaching and learning. Through descriptive or prescriptive analyses, foundations become the basis for identifying the domains to be covered. *See* Criterion, Model.

**Frequency of evaluation** – how often an evaluation is conducted (e.g., every year for a tenured teacher, three times a year for a non-tenured teacher).

**Generalizability** – the appropriateness of extending a process, program, policy, or result across different contexts. *See* Context (teaching), Situational specificity, Transportability, Validity, Validity generalization study.

**Goal** – a statement of intent or an end that a person or a group strives to attain. A goal tends to be more general than an objective. *See* Objectives, Performance goal.

**Grievance** – a claim by a teacher who has been evaluated that the results misrepresent the teacher's level of performance. A grievance may lead to the filing of an appeal and possibly to a formal hearing. *See* Appeal process, Hearing.

**Halo effect** – a type of rater effect in which an evaluator's rating on some dimension affects (in the same direction) his or her ratings on another dimension, or overall. *See* First-impression effect, Rater effect.

**Hearing** – an opportunity for a teacher facing dismissal to appeal the decision by presenting arguments, proofs, evidence, and testimony by others. The process includes several steps such as discovery, direct examination, cross and re-direct examinations, closing arguments, deliberation, and the issuance of a written ruling. *See* Appeal process, Collective bargaining agreement, Dismissal, Grievance.

**High-inference** – types of judgments, decisions and conclusions that are based on complex inductive reasoning by the judge or evaluator. Examples include: inferring to the extent of teacher responsibility for students being off task, in a classroom afflicted by a high external noise level. These judgments tend to be less valid than low-inference judgments, and hence systems of teacher evaluation try to minimize them, to the extent compatible with validity. *See* Inference, Low-inference.

**Holistic (or wholistic) scoring** – the assignment of a single score that reflects an overall impression of performance on a measure. Scores are defined by prescribed descriptors of levels of performance or scoring rubrics. *See* Analytic scoring, Scoring, Scoring rubric.

**Incentive pay** – the allocation of special payments or salary increments to a teacher who does different types of work or assumes additional responsibilities (e.g., coaching an athletic team, being a mentor teacher, teaching a particularly challenging group of students or a difficult course). *See* Longevity pay, Merit pay.

**Incompetence** – the intentional or unintentional failure to perform the duties and professional responsibilities of the teaching job in a minimally acceptable manner as specified by the employing district. Incompetence usually results in remediation, reassignment, or dismissal. *See* Dance of the lemons, Dismissal, Remediation.

**Inconsistency** – a type of rater effect in which an assessor or an evaluator tends to rate or to interpret similar data and information in different ways. *See* Rater effect.

**Indicator** – *See* Criterion.

**Inference** – a conclusion explicitly supported by data, evidence and information gathered as part of the teacher evaluation process. *See* Data, Evidence, High-inference, Information, Low-inference.

**Informal** – the conducting of an evaluation activity without a prescribed plan or structure, or with little or no advance notice. *See* Formal.

**Information** – the knowledge about the attributes and performance of a teacher, based on assessments, documentation, interviews, observation, and other data sources used in the evaluation process.

**Informed consent** – agreement between concerned parties about the data-gathering process, and/or the disclosure, reporting, and/or use of data about a teacher's achievements, and/or the results of a teacher's evaluation. *See* Reporting.

**Input variables** – the activities, materials, and teacher behaviors designed to improve student learning and behavior. Examples include lesson plans, teacher knowledge of the topic being taught, preparation of equipment, and teacher awareness of student misconceptions. *See* Outcome variables.

**Instrument** – a device used to collect data, information and evidence. These can include tests, questionnaires, interview schedules, checklists, rating scales, and observation forms. *See* Measure (*noun*).

**Intern** – a new teacher who receives support from a mentor teacher as part of a professional development process. *See* Mentor teacher.

**Interview** – a process which in part comprises a series of verbally-delivered questions designed to elicit responses concerning attitudes, information, interests, knowledge, quick-response skills, and opinions. The interview has other functions, e.g., when used for selection, it also serves to answer questions from the candidate, and to make the case—to an outstanding candidate—for accepting an offer of the position for which the interview is being conducted, rather than any other offer which may also be received.

**Interviewer** – an assessor who takes part in the interview, either in a face-to-face setting or by telephone; someone who should be making a record of the responses.

**Job analysis** – a technique for studying a teaching job in terms of the knowledge, skills and abilities (KSAs) needed, as well as the functions and tasks that are performed as part of the job.

**Job assignment** – the designated position of a teacher, including the grade levels and the students to be taught, the curricular areas to be covered, the work location, the duration of employment, and non-instructional responsibilities.

**Job description** – a summary of the qualifications, duties, responsibilities, and working conditions associated with a specific job; important in legal matters. It tends to focus on the *distinctive* aspects of the job, so it does not mention many of the duties of the teacher that are understood to be part of all teachers' jobs.

**Joint Committee on Standards for Educational Evaluation** – a group representing most of the relevant professional associations, that developed 21 standards for use in assessing personnel evaluation systems. The standards are grouped under the headings of accuracy, feasibility, propriety, and utility. Listed in references.

**Judge** – a person who makes judgments for use in evaluating the teacher. *See* Assessor, Evaluator, Judgment.

**Judgment** – an appraisal, decision, or opinion about the performance level of a teacher with respect to the knowledge, skill, ability, or behavior being assessed.

**Knowledge** – the sum of the information and experience the teacher has acquired or learned and is able to recall or use.

**KSAs** – knowledge, skills, and abilities. *See* Job analysis.

**Learning environment** – the setting in which student instruction occurs. *See* Context (teaching).

**Leniency** – a type of rater effect in which an assessor or an evaluator tends to rate a teacher too high or to judge the performance level as better than it actually is. *See* Rater effect, Stringency.

**Lesson** – the content that is to be taught or the activity that is to be done during a specific period of instructional time.

**License** – the approval by a governmental agency, usually at the state level, for an individual to be a teacher in the designated state, grade levels, subject areas and specialties. Licensure indicates that the candidate has met a minimum level of requirements, designed to ensure the protection of student welfare, public health, and safety interests. *See* Certification, Credentialing.

**Log (teacher)** – a journal or diary, maintained by the teacher, assessor, supervisor, or administrator, that includes such topics as decisions, plans, activities, results, changes and reflections. The log can serve as a source of information for an evaluation or be included as part of a portfolio. *See* Portfolio, Reflection.

**Longevity pay** – salary increases that are based solely on accrued time of service. *See* Incentive pay, Merit pay.

**Low-inference** – the types of judgments, decisions and conclusions that require a low degree of subjectivity on the part of the judge or evaluator (e.g., "teacher emphasizes recall questions," based on a count of questions asked of students during a classroom observation and the questions on the teacher's unit test). *See* High-inference, Inference, Objective, Subjective.

**MBO (Management by Objectives)-based evaluation** – a teacher evaluation approach based on pre-specified objectives set for or in collaboration with the teacher. *See* Foundation.

**Master teacher** – a teacher who has been identified as exhibiting superior performance and expertise.

**Measure** *(noun)* – an instrument or device that provides data on the quantity or quality of that aspect of teaching performance being evaluated. *See* Instrument.

**Measure** *(verb)* – to determine the degree of quality or quantity of that aspect of teaching being evaluated.

**Measurement-based evaluation** – a teacher evaluation approach that is based on the use of methods which are not dependent on the expertise of the evaluator. *See* Evaluator, Foundation.

**Mentor or mentor teacher** – an experienced, often specially trained, teacher who works with new teachers, interns, or regular teachers in a professional improvement program. Mentors serve as resources, coaches, advisors and confidants to other teachers and are involved in formative evaluation activities as well as in the development and implementation of the plan of assistance. *See* Coaching, Intern, Master teacher, Modeling, Plan of assistance.

**Merit** – the professional competence of a teacher. Merit can include such factors as breadth and depth of knowledge of a subject area, specialized training completed, ability to work with different types of students, and fluency in a second language. *See* Competence, Value, Worth.

**Merit pay** – the allocation of special payments or salary increments to a teacher, based on some form of evaluation of teaching performance that supposedly demonstrates that teacher's superior performance. *See* Incentive pay, Longevity pay, Meritorious performance.

**Meritorious performance** – a level of performance that well exceeds the standard for minimally acceptable and that may be worthy of professional recognition, career ladder advancement, or merit pay. *See* Merit pay, Minimally acceptable.

**Method of evaluation** – the approach used to conduct the evaluation (e.g., the use of formal classroom observations followed by an interview with the supervisor and an oral examination by a team of peers).

**Minimally acceptable** – a performance level that meets the minimum standards for a competency as defined by its criterion. *See* Competency, Standard.

**Model** – an example of a coherent method, approach, procedure or strategy of teaching or of teacher evaluation, as defined by its key or unique assumptions, propositions, attributes, supportive theory, research and practical precedent, or foundation, and which implicitly defines accomplished or good teaching. *See* Criterion, Foundation.

**Modeling** – the use of exemplary teachers and mentors to demonstrate practices of good teaching to other teachers, for the purpose of improvement or of repertoire expansion. *See* Exemplary teacher, Mentor teacher, Self-assessment.

**Monitoring** – checking on a process or person to verify that progress is being made, required or suggested activities are occurring, evaluation procedures are being implemented, prior information is still applicable, earlier decisions can still be justified, and/or standards are being met.

**Multiple-cutoff model** – The simple conjunctive model, requiring a certain level of performance on several dimensions or components. *See* Compensatory model, Conjunctive model, Multiple-hurdle model.

**Multiple-hurdle model** – A sequential conjunctive model; several 'tests' (which might be interviews or simulations or file searches) are administered sequentially, each of which must be passed. Has the advantage over the plain multiple-cutoff approach that the number of test administrations and hence the cost goes down across the series, especially if the most severe tests are done first. *See* Multiple-cutoff model.

**Notes** – descriptive information about the context within which the teacher is being evaluated and/or about the teaching itself. For a particular observation, this might include the number and types of students present, a list of materials being used in the lesson, a description of the classroom arrangement, and information on events that occur during the observation.

**Objective** *(adjective)* – a characteristic of an assessment, observation or conclusion that *either* (1) implies minimal impact of bias and subjectivity (i.e., dispassionate); *or* (2) refers to phenomena which can be inspected or inferred by others (i.e., intersubjective). Sense 1 is evaluative, sense 2 is not. Both often apply, but a person's report that he or she is in pain can be entirely objective in Sense 1, although it is not objective in Sense 2, so it is fallacious to identify the two senses. Both types of claim can be confirmed, although in different ways. *See* Low-inference, Subjective.

**Objectives** – the pre-specified intended outcomes of a program, process, or policy. In the case of education, these are usually in the form of learning and behavioral objectives for students. Professional development objectives may also be part of the teacher evaluation process. Objectives are more specific and closer to the observation level than goals. *See* Goals.

**Observation** – one of several methods used to collect data about a teacher's performance. It may also cover student behavior, the teaching context and the learning environment. Observing should include the recording of evidence and notes while watching the teacher. Observations typically occur in the teacher's own classroom, but they may also occur in other settings (e.g., playground, staff meeting, parent-teacher conference) or may be based on audiotapes or videotapes. *See* Announced observation/visit, Behavior, Context (teaching), Learning environment, Observer, Performance (teacher).

**Observer** – the person who collects evidence and notes about what he/she is observing either in a classroom or another setting. The observer may or may not be an evaluator. *See* Assessor, Evaluator, Observation.

**Orientation** – the acquainting of some or all stakeholders with the teacher evaluation policies, procedures, and processes, in order to promote understanding and improve the quality of the evaluation. *See* Stakeholders.

**Outcome variables** – the results or products of teaching. Examples include student mastery of skills, completion of courses, teacher-developed instructional materials, student projects, and student performance on standardized tests. *See* Input variables, Student learning outcomes.

**Passing score** – a score or level of performance that represents the difference between those teachers whose level of performance is minimally acceptable and those whose performance is not acceptable. Teachers below the "passing score" (also called cut score or critical score) may be given a plan of assistance, mentoring, a job reassignment, or even be dismissed from teaching, depending on the level of performance and the merit and worth of that teacher. *See* Cutting score, Dismissal, Merit, Plan of assistance, Worth.

**Pattern** – a series of similar behaviors or common attributes over a period of time or across different settings or contexts.

**Pedagogy** – the art and science of teaching. Some pedagogical skills apply across most teaching situations, others only to specific subject areas (pedagogical content knowledge).

**Peer review** – the evaluation of a teacher by other teachers, usually done to provide feedback to the evaluee for purposes of professional development and improvement, or to provide an evaluation of subject-matter or context-related expertise not feasible by others involved in the evaluation process. Sometimes used as part of a system of summative teacher evaluation. *See* Coaching.

**Performance (teacher)** – the activities of a teacher on the job. Performance is a function of the teacher's competence and abilities as well as of the context within which the teacher works. *See* Ability, Competence, Context (teaching).

**Performance assessment (performance evaluation)** – an evaluation approach in which teacher merit, worth, or value is determined by rating the teacher's performance and products, according to pre-specified standards or scoring rubrics. *See* Assessor, Attribute, Merit, Performance (teacher), Standard, Value, Worth.

**Performance goal** – a specific statement of what is to be accomplished by the teacher (e.g., relating to growth in knowledge, development of a skill, changes in practice), how the goal will be met (e.g., activities, resources), when the goal will be met, and how achievement of the goal can be assessed or determined.

**Performance standard** – *See* Standard.

**Permanent teacher** – a teacher who has a permanent contract with the employing school district or educational agency. *See* Tenure, Tenured teacher.

**Personnel evaluation** – the systematic determination of the merit, value and worth of the job-related performance of an employee. *See* Merit, Value, Worth.

**Personnel evaluation standards** – *See* Standards.

**Plan of assistance** – a strategy for initial or remedial professional development and growth designed to address a teacher's deficiencies in meeting designated performance standards, based on the results of an evaluation. The plan of assistance should indicate goals and objectives for improvement, an action plan for

improvement, what staff and resources are available, the time line for development activities, benchmarks for ensuring that professional growth is occurring, and measures for verifying achievement of the goals and objectives. *See* Contract, Diagnosis, Performance goal, Remediation.

**Policy (teacher evaluation)** – a set of mandates, rules and guidelines issued by a governmental or administrative agency regarding the purpose of teacher evaluation and the manner in which it should be conducted. *See* Practice (evaluation), Procedures (evaluation).

**Portfolio** – a collection of documents concerning a teacher's performance (e.g., testimonials, student learning outcome reports, self-evaluation, samples of students' work) and of products produced by the teacher (e.g., innovative lesson plans, a critique of a textbook chapter, action research results, a videotape of a lesson, some recent teacher-made unit tests).

**Practice (evaluation)** – the manner in which evaluations are actually conducted, whether or not the practice is in accordance with the policy and/or follows the procedures. *See* Policy (teacher evaluation), Procedures (evaluation).

**Predictive validity** – *See* Validity.

**Preponderance** – the emphasis or weight given to an attribute of a teacher, or to the evidence collected about an attribute. This weight may be based on such considerations as the quantity and frequency of occurrence, its importance to the job of teaching, and its potential impact on the students. As used in teacher evaluation, the term does not connote dominant weight. *See* Attribute.

**Probationary teacher** – a non-tenured teacher who is usually relatively inexperienced (three years or less of teaching experience). *See* Tenure, Tenured teacher.

**Procedures (evaluation)** – the directions for implementing all aspects of the evaluation process in accordance with the rules and guidelines given in a district's policy. Procedures specify how the evaluation is to be conducted, designated time lines, persons responsible, forms to be used, documentation to be provided, the analysis plan, and the steps to be followed. It is often argued that a procedures manual should be in the library in every school or a copy given to each teacher. *See* Data collection procedures, Policy (teacher evaluation), Practice (evaluation).

**Process-product research** – *See* Effective teaching.

**Productivity (teacher)** – the extent of the accomplishment of the primary functions of teaching, including the promotion of increased student learning and of improved student behavior within the teaching context.

**Professional development** – a process designed to improve specific professional competencies or the overall competence of a teacher. *See* Competence, Competency, Formative teacher evaluation, Plan of assistance, Remediation.

**Professionalism (teacher)** – the view—and the associated reform movement —that teaching should be promoted as a profession with its own knowledge base, licensure structure, standards for practice, and professional functions.

**Proficiency** – expertise in a knowledge area or mastery of a skill.

**Profile** – a representation of a teacher's performance on a number of attributes, measures or dimensions.

**Propriety** – the quality of being conducted in a proper, legal, and ethical manner with due regard to the welfare of all involved in and affected by the results of the evaluation. One of the general standards in the Personnel Evaluation Standards (listed in the references).

**Psychomotor domain** – the range of locomotor behaviors needed to explore the environment and perform tasks as well as the sensorimotor activities that are essential to learning and communication. *See* Affective domain, Cognitive domain.

**Qualitative information** – facts and evidence recorded in narrative, graphic, audio, or visual form, but not expressed numerically. *See* Quantitative information.

**Quality check** – the process of verifying both the accuracy of specific data and information, and the appropriateness of the techniques used to collect, arrive at and analyze the data and information (e.g., trained observer used, proper timing given for a performance task, right scoring key used for a test, correct formula used to weight various scores). *See* Accuracy, Audit, Verification.

**Quantitative information** – the facts and evidence that are recorded in numerical form, in tables or scientific graphs, or which could in principle be so recorded. *See* Qualitative information.

**Questionnaire** – an instrument consisting of a series of queries and statements that is used to collect data, reactions, and information from a teacher concerning such factors as educational background, goals and objectives, instructional plans, teaching context, attitudes, opinions, and professional activities; and from others (e.g., students, parents) concerning the teacher's performance.

**Rater effect** – the tendency of an assessor or an evaluator to rate a teacher's performance at a level that does not accurately or consistently reflect the level of his or her performance. There are several types of rater effect, all of which are possible sources of systematic error of measurement. *See* Assessor, Bias, Central tendency effect, Contamination, Contrast effect, Error of measurement, Evaluator, First-impression effect, Halo effect, Inconsistency, Leniency, Reliability, Similar-to-me effect, Stringency.

**Rating** – an estimation, judgment, or calculation of the magnitude or degree of some attribute of teaching, using a numerical or descriptive continuum.

**Recommendations** – a set of suggestions derived from the teacher evaluation results. For formative teacher evaluation, they may include a list of professional development activities and a plan of assistance. For summative teacher evaluation, they may consist of personnel actions such as tenure, dismissal/termination, reassignment/transfer, contract renewal, or promotion.

**Record** *(noun)* – the written or taped data, evidence, judgments, notes, recommendations and other statements for use in the teacher evaluation process. *See* Documentation.

**Record** *(verb)* – to register and store data and other information. *See* Record.

**Reflection** – the process by which a teacher reviews his or her past performance as a means of improving future performance; *or* the process by which a teacher decides on tactics in the classroom in the light of all relevant factors present. *See* Log (teacher), Self-assessment.

**Relevance (domain)** – the extent to which the domains and indicators covered by a teacher evaluation system apply to a teacher's professional functions in terms of both importance and frequency.

**Reliability** – the degree to which an assessment, measure, or instrument consistently measures an attribute. There are several types of reliability, namely: (1) the degree to which the measure yields consistent results over different administrations with the same teacher performing at the same level, by the same assessor (intra-rater), (2) the degree to which the measure yields similar results for the same teacher at the same time with more than one assessor (inter-rater), (3) the degree to which individual observations or items consistently measure the same attribute (internal consistency), and (4) the degree to which the measure produces consistent results over several administrations assessing the same attribute of a teacher (test-retest). *See* Error of measurement, Reliability coefficient.

**Reliability coefficient** – a measure of the reliability of an instrument, ranging from 0.00 to +1.00. The closer the reliability coefficient is to +1.00, the more reliable is the measure. *See* Reliability.

**Remediation** – those techniques or strategies designed to improve a teacher's performance with respect to general deficiencies or specific areas of weakness. *See* Clinical supervision, Coaching, Diagnosis, Feedback, Formative teacher evaluation, Incompetence, Modeling, Plan of assistance, Professional development, Reflection, Self-assessment, Teacher improvement.

**Reporting** – the process of communicating results and recommendations to designated individuals or groups. When reporting to the teacher, this would be considered part of feedback. *See* Feedback, Informed consent.

**Research-based evaluation** – a teacher evaluation approach that is based on the use of at least some "empirically-validated" criteria or indicators of competence derived from research studies of effective teaching practices. *See* Effective teaching, Foundation.

**Responses** – the answers to test, interview, or questionnaire items.

**Responsibility** – that which a person is expected and obligated to do and for which he/she is accountable. *See* Duty.

**Review** – to examine again or to look at thoroughly in order to make a judgment, come to a conclusion, or evaluate.

**Sampling of performance** – the selection of an array of performances and settings to be evaluated so that they are wide enough in range and large enough in number to cover the scope of the performances addressed by the evaluation system. That is, they need to be comprehensive and representative, in order to support valid inferences about merit. *See* Error of measurement.

**Schedule** – the designated dates and times for various activities related to the evaluation (e.g., when the classroom will be observed, when the portfolio is to be submitted, when the teacher will be interviewed). The schedule, which is part of the procedures, should also indicate the location of such activities and the people who will be involved; it should be included in the procedural manual. See Procedures.

**Scope of content** – the extent of coverage, by an instrument or process, of all knowledge, skills, abilities, and behaviors to be measured, in terms of both breadth and depth of coverage.

**Score** *(noun)* – the number of points earned on an assessment or the degree of success on a measure of teacher attributes or performance.

**Scorer** – an assessor who awards a score on one part or all of an assessment; may also be the evaluator. *See* Assessor, Evaluator, Scoring.

**Scoring** – the summarizing of assessment results in terms of the degree to which the teacher meets the standards for that assessment. Scores are usually expressed in numerical terms, but in some cases in descriptive terms or graphically. *See* Analytic scoring, Holistic scoring, Scorer.

**Scoring rubric** – a set of rules, guidelines, or prescribed descriptors for use in quantifying measures of teacher attributes and performance. *See* Analytic scoring, Holistic scoring.

**Scripting** – the writing of evidence and notes throughout an evaluation activity (e.g., interview, classroom observation) about what is happening. Scripting does not include making high-inference judgments or interpretations of the evidence and notes. Scripting may be continuous throughout the activity or intermittent at pre-specified time intervals (e.g., 3 minutes on, 2 minutes off).

**Self-assessment** – the process of judging one's own teaching performance for the purpose of self improvement. A teacher may use such techniques as self-viewing on a videotape, observing and modeling exemplary teachers (and perhaps asking them to rate the implementation), self-rating forms, open-ended self-reports, portfolios, or self-study handbooks and materials. *See* Reflection, Self-evaluation, Self-study materials.

**Self-evaluation** – the process of reviewing one's own behavior and student learning outcomes for the purpose of monitoring and changing one's own teaching performance. *See* Reflection, Self-assessment, Student learning outcomes.

**Self-study materials** – the programs designed for use by a teacher in assessing his or her own teaching behaviors. *See* Self-assessment.

**Sensitivity** – awareness and understanding of other people's feelings, attitudes, social and cultural backgrounds, ethnic traditions and customs, languages, interests, rights, and needs.

**Setting** – the temporal and physical environment of an event or activity. *See* Context (teaching), Learning environment.

**Similar-to-me effect** – a type of rater error in which an assessor or an evaluator judges higher or more favorably those people seen as similar to him/herself. *See* Rater effect.

**Simulation** – a structured exercise or a re-creation of a typical job task or situation to assess how well a teacher can implement such a task or perform in such a situation (e.g., asking a science teacher to prepare slides for biology, having a music teacher listen to a tape and show how he/she would conduct that piece of music using the musical score in front of him/her, asking a fourth-grade teacher to develop and present a learning activity on interpreting map symbols). Suffers the usual drawbacks of high-stakes assessment—at best, it only tests peak per-

formance, at worst it is invalid because of stress impairment .

**Situational specificity** – the extent to which it is inappropriate to extend information on and conclusions about a teacher's performance in a certain context to other contexts. *See* Context (teaching), Generalizability.

**Skill** – the ability to use knowledge in a practical manner or simply to do something. *See* Ability, Knowledge.

**Specifications (assessment)** – a delineation of the major attributes of an assessment to be developed, including breadth and depth of content to be covered, level of difficulty, format of the assessment materials, scoring procedures, and numbers and types of items.

**Stakeholders** – those individuals who will be significantly affected by the results of teacher evaluation. These include not only teachers and principals, but also other teachers, school and district staff, students, parents, school board members, future employers and fellow citizens, taxpayers, and community members. They are part of the audience for evaluations of teaching, but have more at stake than just potential interest. Their interests are not always paramount, since they may include improper interests (e.g., prevention of serious questioning of authority, religion, or creationism), hence serving their wishes cannot be done without examination of the justification of their preferences. *See* Audience.

**Standard** – the level of performance on the criterion being assessed that is considered satisfactory (or excellent, etc.) in terms of the purpose of the evaluation. Developmental standards specify improvement levels to be attained and may be used for professional development and self-assessment. Minimum standards designate the level below which performance is not acceptable and are used for such purposes as licensure and job assignments or termination. Desired performance standards reflect what is regarded as effective or accomplished teaching, and typically are used for such purposes as promotions, awards and certification. *See* Criterion.

**Standards** – shorthand for the ***Personnel Evaluation Standards.*** *See* Joint Committee.

**Standard setting** – the determination of the teaching performance level considered acceptable (excellent, etc.) in terms of the purpose of the evaluation. Standards are usually determined using empirical or judgmental techniques or a combination of these. *See* Standard.

**Standardization** – the use of consistent procedures for administering, scoring, reviewing, interpreting, and reporting the results of the teacher evaluation.

**Stringency** – a type of rater effect in which an assessor or an evaluator tends to rate a teacher too low or to judge the performance level as poorer than it actually is. Sometimes referred to as "negative leniency." *See* Leniency, Rater effect.

**Student learning outcomes** – measures of student achievement of knowledge and skills, and other educational outcomes such as improved student attitudes and behaviors, that should have been taught to them by the teacher being evaluated. This term covers acquisition, retention, application, transfer, and adaptability of knowledge and skills. *See* Outcome variables.

**Style-based evaluation** – a teacher evaluation approach using criteria that specify in what way or how teachers are to perform their job (e.g., by use of high eye contact, high frequency of questioning, cooperative learning, announced lesson objectives) rather than measures of the degree to which teachers get the job done. *See* Foundation.

**Subjective** *(adjective)* – an observation, assessment, or conclusion that is *either* (1) a personal, idiosyncratic, view (e.g., a biased view or 'mere opinion') that does not agree with the judgments of other, equally well-qualified, judges; *or* (2) a reference to the private experiences or feelings of the person who makes the claim. In the first but not the second sense, subjective is an evaluative term about the quality of evidence or conclusions, so it is incorrect to identify the two senses. *See* Audit, Judgment, Objective, Verification.

**Summary** – a concise report encompassing the major results of an evaluation of a teacher's performance.

**Summative teacher evaluation** – an evaluation conducted primarily for the purpose of making personnel decisions about the teacher (e.g., merit pay, reassignment, promotion, dismissal, tenure). Usually done by a senior administrator such as a principal, rather than by another teacher or a supervisor (where that position is not occupied by the principal). *See* Formative teacher evaluation.

**Supervisor** – the person responsible for overseeing the work of a teacher and for ensuring that the teacher performs his or her duties and professional responsibilities. The job normally requires substantial formative evaluation activities but typically not full-scale summative evaluation, although the same person may do both. *See* Administrator.

**Synthesis** – the combining of data and information from multiple sources, or of ratings and judgments on separate dimensions, for use in arriving at (typically) an evaluative conclusion or result. Often necessary for summative evaluation, sometimes unnecessary for formative evaluation. *See* Analysis, Conjunctive, Disjunctive, Formative, Review, Summative.

**System (evaluation)** – a complete approach to the evaluation of teachers including its purpose, the rules and regulations that apply, the target group to be evaluated, the domains to be covered, the procedures and methods to be employed, the instruments to be used, the persons to be involved, and the types of reports and feedback to be provided.

**Teacher effectiveness** – the degree to which a teacher has a positive impact on student learning, behavior and attitudes. *See* Effective teaching.

**Teacher evaluation** – the process of determining the merits, values and worth of a teacher, based on an assessment of the teacher's knowledge, skills, abilities and behavior, and on the results of his or her teaching. *See* Assessment, Evaluation.

**Teacher improvement** – the accomplishment of goals and objectives for professional development, growth in knowledge, acquisition of skills, and changes in practices. *See* Plan of assistance, Remediation.

**Teacher norms** – the expectations of teacher behavior in a given context or setting that are usually learned during student teaching or in the first year of teaching in a new school.

**Teachers' rights** – the privileges of teachers guaranteed by law, standard practice, ethics, and under their collective bargaining agreement. They include the rights to form and organize a teachers' association, to select their representatives, and to bargain collectively. *See* Collective bargaining agreement, Collective bargaining unit.

**Temporary teacher** – a teacher who is filling a position for a specified period of time, whether to fill a vacancy on a short-term basis (one year or less) or for an experimental teaching position with a designated ending date. Occasionally used to refer to teachers who have been on the job for many years and are there as long as the district can afford them.

**Tenure** – an employment status conferred upon a teacher by state law or institutional regulation after successful completion of a probationary period. Tenure provides substantial but not complete protection against arbitrary or capricious dismissal and entitles the teacher to due process procedures and other protections that may not be available to the non-tenured teacher. *See* Dismissal, Due process, Permanent teacher, Probationary teacher, Tenured teacher.

**Tenured teacher** – a teacher who has completed a probationary period (usually two to five years in K–12, seven in the post-secondary area) and is now considered a permanent employee of the school district with all applicable benefits and rights as specified in state law, district policy and the collective bargaining agreement. *See* Collective bargaining agreement, Permanent teacher, Probationary teacher, Tenure.

**Test pollution** – the results of practices designed to increase test scores or performance ratings without improving actual performance of the attributes being assessed. A by-product of high-stakes testing of students or teachers. *See* Coaching.

**Theory-based evaluation** – a teacher evaluation approach that is based on certain theories of teaching or of learning. *See* Foundation.

**Time line** – a calendar or list of dates showing the evaluation stages and activities, and indicating the dates by which they should be implemented and be completed.

**Track record** – a summary of past events and accomplishments related to a teacher's performance. In the narrow sense, it refers only to documented teaching performance; but it is often used to include details of further education and training completed, conferences attended, awards received by the teacher and his or her students.

**Transportability** – the appropriateness of extending the use of a policy, instrument, assessment procedure, or evaluation system across different teachers, student groups, subject areas, instructional approaches, learning activities, school settings, states, etc. *See* Generalizability.

**Triangulation** – the attempt to obtain more accurate results by using multiple sources of data about one aspect of performance, multiple methods of collecting data, multiple judges, or multiple interpretations of the same data. *See* Congruence analysis.

**True score** – a hypothetical score that represents an assessment result which is entirely free of error. Sometimes true score is thought of as the average score of

an infinite series of assessments with the same or exactly equivalent instruments, with no practice effect or change in the person being assessed across the series of assessments. *See* Error of measurement.

**Trust** – a common understanding of the purpose and potential of teacher evaluation, and a cooperative spirit between the teacher and the evaluator for maximizing the benefits of doing the evaluation. Trust is related to such factors as confidentiality of communication, careful consideration of the accuracy of evidence from such sources as hearsay or complaints, honesty, openness, sharing, and sincerity on the part of both the teacher and the evaluator. *See* Credible.

**Utility (of a teacher evaluation system)** – the practical value of a teacher evaluation system. Utility involves such factors as time requirements, logistics, money costs and benefits, non-money resources needed, and district and state policies, as well as the technical concerns of validity and reliability. One of the 'Standards.' *See* Administrative feasibility.

**Validation** – the process of assessing the appropriateness, meaningfulness and usefulness of a measurement, an instrument, or a process, and of the inferences made from the results of it. *See* Validity, Validity coefficient.

**Validity** – the extent to which an instrument or process measures the attribute that it is claimed to measure. Several types of validity are commonly referred to in the literature, although they have now been superseded by the more recent classification of the 1985 APA Standards. Those relevant to teacher evaluation are defined below. *See* Generalizability, Validation, Validity coefficient.

**Concurrent** – the relationship of a measure to another simultaneous measure assessing the same attribute.

**Construct or construct-related** – the degree of fit of a measure and its interpretation with its underlying explanatory concepts, theoretical rationales, or foundations.

**Content or content-related** – the extent of congruence between the scope of a content area that an instrument or process claims to cover and what it actually does cover.

**Criterion-related** – the correlation or extent of agreement of the results of an assessment with one or more external variables regarded as direct measures of the attribute being assessed.

**Face** – the extent of apparent acceptability or legitimacy of an instrument or process to teachers, administrators, policymakers, students and parents, the general public and other stakeholders concerned with teacher evaluation and the quality of teaching.

**Predictive** – the relationship of a measure to a future measure or variable assessing a similar or a different (but presumably related) attribute. One type of criterion-related validity.

**Validity coefficient** – a measure of the degree of validity, usually expressed as the correlation between the measure in question and another measure or a variable (a criterion measure). The range is from -1.00 to +1.00, with zero indicating no relationship; values further from zero indicate a stronger (positive or negative) relationship between the measure in question and the other measure or

variable. *See* Validation, Validity.

**Validity generalization study** – an examination of the degree of generalizability or transportability of a policy, instrument, process, procedure, or evaluation system, or of the situational specificity of evaluation results for an individual or an evaluand. *See* Evaluation, Generalizability, Situational specificity, Transportability.

**Value** – an estimation or a measure of the merit and/or worth of a teacher in terms of the quality of the individual teacher (merit) and of the teacher's potential benefit to the school (worth). *See* Merit, Worth.

**Verification** – the process of checking the accuracy of data and information about the teacher's experience, training, performance, and other attributes. *See* Accuracy, Audit, Quality check.

**Weighted score** – a score adjusted by such factors as the importance of the attribute assessed to teaching performance, or the reliability and validity of the assessment from which the score was derived, or a combination of such factors.

**Wholistic** - alternative spelling for holistic.

**Wisdom of practice** – that which a teacher learns about teaching through experience rather than in a traditional teacher-training program.

**Worth** – the value of a teacher to the school as a *system*. For example, being able to speak Spanish could be of high worth in a school that does not have enough Spanish-speaking teachers, whereas being able to coach basketball may be of low worth to the school if there are already several teachers who can coach basketball. The teachers who know Spanish or are able to coach basketball can still be assessed for merit, in addition to their worth, and may or may not have worth at another school or at another time in the same school. *See* Merit, Value.

# RESOURCES

Listed below are several national associations and agencies that provide materials and services related to the evaluation of teachers. Additional resources include state educational agencies; state teacher licensing boards; national subject matter associations for teachers; state and regional associations for teachers, administrators, and school boards; regional educational laboratories; and schools of education at colleges and universities.

American Association of School Administrators (AASA)
1801 North Moore Street
Arlington, VA 22209
(703) 528-0700

American Federation of Teachers (AFT)
555 New Jersey Avenue, N.W.
Washington, DC 20001
(202) 879-4400

Association for Supervision and Curriculum Development (ASCD)
1250 North Pitt Street
Alexandria, VA 22314-1403
(703) 549-9110

Center for Research on Educational Accountability and Teacher Evaluation (CREATE)
The Evaluation Center
Western Michigan University
Kalamazoo, MI 49008-5178
(616) 387-5895

Education Commission of the States (ECS)
707 17th Street, Suite 2700
Denver, CO 80202-3427
(303) 299-3600

Educational Research Service (ERS)
2000 Clarendon Blvd.
Arlington, VA 22201
(703) 243-2100

ERIC Clearinghouse on Educational Management
College of Education
University of Oregon
1787 Agate Street
Eugene, OR 97403
(503) 346-5044

ERIC Clearinghouse on Teacher Education
American Association of Colleges for Teacher Education
One Dupont Circle, N.W., Suite 610
Washington, DC 20036
(202) 293-2450

ERIC Clearinghouse on Tests, Measurement, and Evaluation
American Institutes for Research
3333 K Street N.W., Suite 300
Washington, DC 20007
(202) 342-5060

National Association of Elementary School Principals (NAESP)
1615 Duke Street
Alexandria, VA 22314
(703) 684-3345

National Association of Secondary School Principals (NASSP)
1904 Association Drive
Reston, VA 22091
(703) 860-0200

National Education Association (NEA)
1201 16th Street, N.W.
Washington, DC 20036
(202) 822-7750

National Organization on Legal Problems in Education (NOLPE)
3601 S.W. 29th Street, Suite 223
Topeka, KS 66614
(913) 273-3550

National School Boards Association (NSBA)
1680 Duke Street
Alexandria, VA 22314
(703) 838-6782

Phi Delta Kappa, Inc. (PDK)
Eighth and Union
P.O. Box 789
Bloomington, IN 47402
(812) 339-1156

The RAND Corporation
1700 Main Street
P. O. Box 2138
Santa Monica, CA 90406-2138
(213) 393-0411

# REFERENCES


Bickers, Patrick M. (1988). *Teacher evaluation: Practices and procedures.* Arlington, VA: Educational Research Service.

Bridges, Edwin M., with the assistance of Barry Groves. (1990). *Managing the incompetent teacher* (2nd edition.). Eugene, OR: University of Oregon, College of Education, ERIC Clearinghouse on Educational Management.

Committee to Develop Standards for Educational and Psychological Testing of the American Educational Research Association, the American Psychological Association, & the National Council on Measurement in Education. (1985). *Standards for educational and psychological testing.* Washington, DC: American Psychological Association.

Duke, Daniel L., & Stiggins, Richard J. (1986). *Teacher evaluation: Five keys to growth.* Washington, DC: American Association of School Administrators; National Association of Elementary School Principals; National Association of Secondary School Principals; National Education Association.

Joint Committee on Standards for Educational Evaluation. (1988). *The personnel evaluation standards: How to assess systems for evaluating educators.* Newbury Park, CA: Sage Publications.

Lewis, Anne C. (1982). *Evaluating educational personnel.* Arlington, VA: American Association of School Administrators.

Linn, Robert L. (Ed.) (1989). *Educational measurement* (3rd edition.). New York, NY: American Council on Education and Macmillan Publishing Company.

Louisiana Department of Education. (1990, July). *Implementation guide for the Louisiana Teaching Internship Program and Louisiana Teacher Evaluation Program* (Bulletin 1877). Baton Rouge, LA: Louisiana Department of Education, Office of Research and Development, Bureau of Professional Accountability.

Millman, Jason, & Darling-Hammond, Linda. (Eds.) (1990). *The new handbook of teacher evaluation: Assessing elementary and secondary school teachers.* Newbury Park, CA: Sage Publications.

Mitchell, James V., Jr., Wise, Steven L., & Plake, Barbara S. (Eds.) (1990). *Assessment of teaching: Purposes, practices, and implications for the profession.* Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.

Pendleton Public Schools. (1990). *Evaluation handbook for the improvement of instruction.* Pendleton, OR: Pendleton Public School District 16R.

Poggio, John P., Burry, Judith A., & Glasnapp, Douglas R. (1988-1989). *Kansas Internship Assessment Inventory: User's manual.* Lawrence, KS: Center for Educational Testing and Evaluation, University of Kansas.

Schubert, William H. (1986). *Curriculum: Perspective, paradigm, and possibilities.* New York, NY: Macmillan Publishing Company.

Scriven, Michael. (1991). *The evaluation thesaurus* (4th edition.). Newbury Park, CA: Sage Publications.

Stiggins, Richard J., & Duke, Daniel. (1988). *The case for commitment to teacher growth: Research on teacher evaluation.* Albany, NY: State University of New York Press.

Stronge, James H., & Helm, Virginia M. (1991). *Evaluating professional support personnel in education.* Newbury Park, CA: Sage Publications.

Tennessee State Department of Education. (1989). *Career ladder teacher orientation manual.* Nashville, TN: The Author.

Walberg, Herbert J., & Haertel, Geneva D. (Eds.) (1991). *The international encyclopedia of educational evaluation.* Oxford, England: Pergamon Press.

Wheeler, Patricia, & Page, Janelle. (1990). *Development of a science laboratory assessment for new teachers, grades K-12: California New Teacher Project.* Mountain View, CA: RMC Research Corporation.

Wittrock, Merlin C. (Ed.) (1986). *Handbook of research on teaching* (3rd edition.). New York, NY: Macmillan Publishing Company.